Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

RALF BORNDÖRFER MARIKA NEUMANN MARC E. PFETSCH

The Steiner Connectivity Problem

Supported by the DFG Research Center MATHEON "Mathematics for key technologies".

ZIB-Report 09-07 (April 2010)

The Steiner Connectivity Problem[§]

Ralf Borndörfer* Marika Neumann* Marc E. Pfetsch**

Abstract

The Steiner connectivity problem is a generalization of the Steiner tree problem. It consists in finding a minimum cost set of simple paths to connect a subset of nodes in an undirected graph. We show that polyhedral and algorithmic results on the Steiner tree problem carry over to the Steiner connectivity problem; namely, the Steiner cut and the Steiner partition inequalities, as well as the associated polynomial time separation algorithms, can be generalized. Similar to the Steiner tree case, a directed formulation, which is stronger than the natural undirected one, plays a central role.

1 Introduction

The *Steiner connectivity problem* (SCP) can be described as follows. We are given an undirected graph $G = (V, E)$, a set of *terminal nodes* $T \subseteq V$, and a set of (simple) *paths* $\mathcal{P}$ in G. The paths have nonnegative costs $c \in \mathbb{R}^{\mathcal{P}}_{+}$. The problem is to find a set of paths $\mathcal{P}' \subseteq \mathcal{P}$ of minimal cost $\sum_{p \in \mathcal{P}'} c_p$ that *connect the terminals*, i.e., such that for each pair of distinct terminal nodes $t_1, t_2 \in T$ there exists a path from t_1 to t_2 in G that is completely covered by paths of $\mathcal{P}'$. We can assume w.l.o.g. that every edge is covered by a path, i.e., for every $e \in E$ there is a $p \in \mathcal{P}$ such that $e \in p$; in particular, G has no loops. Figure 1 gives an example of a Steiner connectivity problem and a feasible solution.

The SCP is a generalization of the *Steiner tree problem* (STP), in which all paths contain exactly one edge. Similar to the STP with nonnegative costs, see [13, 15, 16] for an overview, there exists always an optimal solution of the SCP that is *minimally connected*, i.e., if we remove any path from the solution, there exist at least two terminals which are not connected. However, in contrast to the STP, an optimal solution of the Steiner connectivity problem does not necessarily form a tree, see again Figure 1.

The SCP is a special case of the *line planning problem*, see [3] and the references therein for a detailed description. The line planning problem can be defined as follows. We are given a public transportation network $G = (V, E)$, a set of (simple) line paths $\mathcal{P}$, and a passenger demand matrix $(d_{uv}) \in \mathbb{N}^{V \times V}$, which gives the number

[§]Supported by the DFG Research Center MATHEON "Mathematics for key technologies"
*Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany; Email `borndoerfer@zib.de`, `marika.neumann@zib.de`

**Institute for Mathematical Optimization, TU Braunschweig, Pockelsstr. 14, 38106 Braunschweig, Germany; Email `m.pfetsch@tu-bs.de`

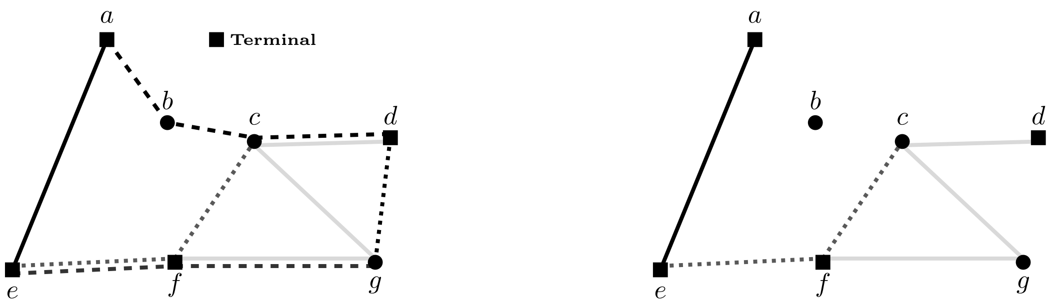

Figure 1: Example of a Steiner connectivity problem. *Left:* A graph with four terminal nodes ($T = \{a, d, e, f\}$) and six paths $\big(\mathcal{P} = \{p_1 = (a, b, c, d), p_2 = (e, f, g), p_3 = (a, e), p_4 = (e, f, c), p_5 = (g, d), p_6 = (f, g, c, d)\}\big)$. *Right:* A feasible solution with three paths ($\mathcal{P}' = \{p_3,\ p_4,\ p_6\}$).

of passengers who want to travel between different stations in the network. The edges of G have nonnegative travel times $\tau \in \mathbb{R}_+^E$, the paths have nonnegative costs $c \in \mathbb{R}_+^{\mathcal{P}}$ and capacities $\kappa \in \mathbb{R}_+^{\mathcal{P}}$. The problem is to find a set of line paths $\mathcal{P}' \subseteq \mathcal{P}$ with associated frequencies $f_p \in \mathbb{R}_+$, $p \in \mathcal{P}'$, and a passenger routing, such that the overall capacities $\sum_{p \in \mathcal{P}', e \in p} f_p \cdot \kappa_p$ on the edges $e \in E$ suffice to transport all passengers. There are two possible objectives: to minimize the travel time, or to minimize the cost of the line paths.

The connection between the line planning problem and the SCP is that the line paths $\mathcal{P}'$ usually connect all stations with positive supply and/or demand. More precisely, let (T, F) be the *demand graph* of the line planning problem, where $T = \{v \in V \mid \sum_u (d_{uv} + d_{vu}) > 0\}$ is the set of nodes with positive supply or demand, and $F = \big\{\{u, v\} \mid d_{uv} + d_{vu} > 0\big\}$ a set of *demand edges*. Then the following holds: If the demand graph is connected, then the set of line paths $\mathcal{P}'$ of a solution of the line planning problem is a solution of the SCP associated with the graph G, terminal set T, and costs c. In other words, if we neglect travel times of the passengers, as well as capacities and frequencies of the lines, the line planning problem with connected demand graph reduces to the Steiner connectivity problem. In this way, the SCP captures the connectivity aspect of the line planning problem. This connection motivates studying the SCP.

A natural question is whether one can transfer structural results and algorithms from the Steiner tree problem to the Steiner connectivity problem. It will turn out that this can indeed be done in many cases. In particular, an important result (see Chopra and Rao [4]) in the STP literature states that the undirected IP formulation of the STP, including all so-called Steiner partition inequalities, is dominated by a certain family of directed formulations. Using this connection, a super class of the Steiner partition inequalities can be separated in polynomial time. We will show that similar results hold for the SCP as well. Namely, a directed formulation of the Steiner connectivity problem, which can be interpreted as an extended formulation, produces a strong relaxation of the Steiner connectivity problem via projection to the original space of variables; see, e.g., Vanderbeck and Wolsey [19] for a discussion of extended formulations. The directed formulation that we use, however, is constructed differently than in the STP case and must be strengthened by so-called flow-balance constraints to obtain an analogous result. Subtle differences also come up in the complexity analysis. For instance, the SCP is also solvable in polynomial time for a fixed number of terminals, but it is NP-hard in the case $T = V$.

The article is structured as follows. It starts with a combinatorial discussion of the Steiner connectivity problem in Section 2. We show that the SCP is equivalent to a suitably constructed directed Steiner tree problem. This relation yields polynomial time algorithms for the SCP in some cases. In Section 3, we give three integer programming formulations for the SCP based on the transformation in Section 2, namely, an undirected cut formulation, a directed cut formulation, and a contracted directed cut formulation. We compare these formulations and their LP relaxations. An analysis of the polytope associated with the undirected cut formulation follows in Section 4. We state necessary and sufficient conditions for the Steiner partition inequalities to be facet defining. We also derive a polynomial time separation algorithm for a super class of the Steiner partition inequalities. This algorithm is based on the directed cut formulation of the SCP. This shows that directed models provide tight formulations for the SCP, similar as for the STP. These theoretical results are illustrated by computations for large scale real-world transportation networks in Section 5.

2 Relation to Directed Steiner Trees & Complexity

We show in this section the equivalence of the SCP and a suitably constructed directed Steiner tree problem. The *directed Steiner tree problem* (DSTP) is the following: Given a directed graph and a set of terminal nodes T, we have to find a minimum cost set B of arcs that connect a *root node* $r \in T$ to each other terminal $t \in T \backslash \{r\}$, i.e., there exists a directed path from r to t in B. If the costs of the arcs are nonnegative, which we assume, there exists a solution that is a directed tree (an arborescence).

Consider an SCP with undirected graph $G = (V, E)$, a set of paths $\mathcal{P}$, terminals $T \subseteq V$, and nonnegative costs $c \in \mathbb{R}^{\mathcal{P}}_{+}$. Define a digraph $D' = (V', A')$, which we call *Steiner connectivity digraph*. Its node set is

$$V' := T \cup \{v_p, w_p \,|\, p \in \mathcal{P}\}.$$

We choose some terminal node $r \in T$ as root node and define the following arcs $a \in A'$ and costs c'_a:

$$\begin{array}{lll} a = (r, v_p), & c'_a := 0, & \forall\, p \in \mathcal{P} \text{ with } r \in p, \\ a = (v_p, w_p), & c'_a := c_p, & \forall\, p \in \mathcal{P}, \\ a = (w_{\tilde{p}}, v_p), & c'_a := 0, & \forall\, p, \tilde{p} \in \mathcal{P},\ p \neq \tilde{p},\ \ p \text{ and } \tilde{p} \text{ have} \\ & & \text{a node } v \in V \text{ in common}, \\ a = (w_p, t), & c'_a := 0, & \forall\, p \in \mathcal{P},\ \forall\, t \in T \backslash \{r\} \text{ with } t \in p. \end{array}$$

Figure 2 illustrates our construction. Note that choosing different root nodes results in different Steiner connectivity digraphs and hence different associated DSTPs. However, we will show in Proposition 2.2 that the solutions of an SCP and any associated DSTP are all equivalent, independent of the choice of the root node. For

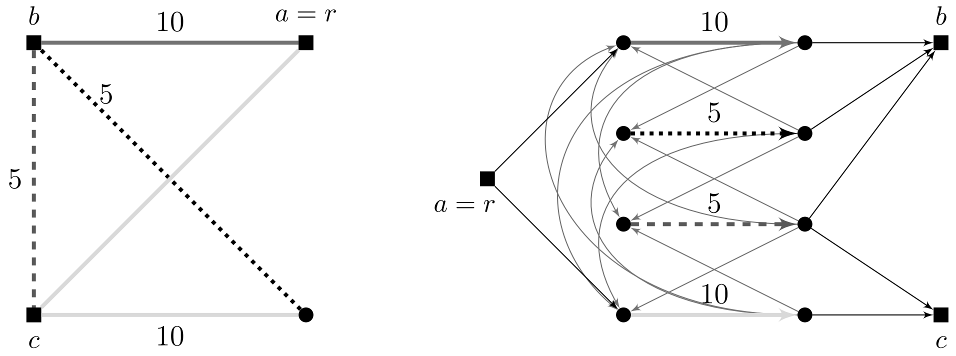

Figure 2: A Steiner connectivity problem and its associated directed Steiner tree problem. *Left:* Graph G with four paths and three terminal nodes. The numbers on the paths indicate costs. *Right:* Associated Steiner connectivity digraph D'. The numbers on the arcs are the costs; the default value is zero.

ease of notation, we will therefore omit the root node whenever the results are independent of r. Polyhedral results can depend on the choice of the root node, see Remark 3.9 below. In such cases we will include the root node in the notation.

A DSTP associated with an SCP has the following properties.

Lemma 2.1. *1. The only arc with target node w_p is (v_p, w_p), for all $p \in \mathcal{P}$.*

2. The only arc with source node v_p is (v_p, w_p), for all $p \in \mathcal{P}$.

3. Each simple directed (r,t)-path, $t \in T \backslash \{r\}$, has the general form $(r, v_{p_1}, w_{p_1}, \ldots, v_{p_k}, w_{p_k}, t)$, $k \geq 1$.

Proposition 2.2. *The following holds for an SCP and an associated DSTP: For each solution of one problem there exists a solution of the other problem with the same objective value. In particular, the optimal objective value of an associated DSTP is independent of the choice of the root node.*

Proof. Assume $\tilde{\mathcal{P}}$ is a solution of SCP. Then let

$$\tilde{A} := A' \setminus \{(v_p, w_p) \,|\, p \notin \tilde{\mathcal{P}}\}.$$

The arcs in $\tilde{A}$ connect the root r with each terminal $t \in T \setminus \{r\}$ via a directed path. Moreover, $\sum_{a \in \tilde{A}} c'_a = \sum_{p \in \tilde{\mathcal{P}}} c'_{v_p w_p} = \sum_{p \in \tilde{\mathcal{P}}} c_p$.

For the converse, assume that $\tilde{A}$ is a solution of the DSTP. We show that

$$\tilde{\mathcal{P}} := \{p \in \mathcal{P} \,|\, (v_p, w_p) \in \tilde{A}\}$$

is a solution of the corresponding SCP with the same cost. To this purpose, consider the root node r and some terminal $t \in T \backslash \{r\}$; these nodes are connected by a simple directed path in D' using only arcs in $\tilde{A}$. Each such path has the form $(r, v_{p_1}, w_{p_1}, \ldots, v_{p_k}, w_{p_k}, t)$, $k \geq 1$ (see Lemma 2.1), with $(v_{p_i}, w_{p_i}) \in \tilde{A}$, $i = 1, \ldots, k$, that is, $p_i \in \tilde{\mathcal{P}}$, $i = 1, \ldots, k$. Due to the construction of D', p_1 contains r, p_i and p_{i+1}, $i = 1, \ldots, k-1$, have at least one node in common, and p_k contains t. Hence, we can find a path from r to t in G that is covered by $p_1, \ldots, p_k \in \tilde{\mathcal{P}}$. Since the paths